

# Scrapy Outcome & Data Export

---



**Eduardo Freitas**

BUSINESS AUTOMATION & DATA CAPTURE SPECIALIST



# Overview



**Feed exporter**

**Refactoring text export code**

**Using an exporter to save data**



# Feed Exporter

---



Allows you to generate a feed with the scraped items, using multiple serialization formats and storage backends.



# Feed Exporter Types



**Serialization formats: JSON, JSON Lines, CSV and XML**



**Storages: Local FS, FTP, S3 and standard output**



**Local FS - the feeds are stored in the local filesystem**



**FTP - The feeds are stored in a FTP server**



**S3 - The feeds are stored on Amazon S3 (requires boto or botocore)**



# Demo



**Using an exporter to save data**



# Summary



**Feed exporter**

**Code refactoring**

**Explore further**

- Broad crawls
- Auto throttling
- Deploying Scrapy to the cloud
- Visual scraping with Portia

